

# Word Saliency and Frequency of Academic Words in Textbooks: A Case Study in the *New Standard College English*

Yan Liqin<sup>1</sup> & Gao Xinlu<sup>1</sup>

<sup>1</sup> Department of Foreign Languages, School of Humanities and Social Sciences, National University of Defense Technology, China

Correspondence: Yan Liqin, School of Humanities and Social Sciences, National University of Defense Technology, Changsha, Hunan, China. E-mail: lqyan305@126.com

Received: November 27, 2013 Accepted: February 7, 2014 Online Published: March 20, 2014

doi:10.5539/ies.v7n4p14 URL: <http://dx.doi.org/10.5539/ies.v7n4p14>

## Abstract

Though textbooks are one of the main vocabulary input resources for domestic college students and core contents of learning and testing (Liu, 2013), few empirical studies are done to evaluate learning opportunities provided by textbooks. This empirical study is designed to analyze what learning opportunity is provided in a currently used series of textbooks of academic words, which in the present study are all from the 570-item Academic Word List (AWL) that Coxhead (2000) produces based on his self-constructed academic corpus. Through the interpretation of the quantitative and qualitative results, it was found that a favorable learning opportunity of academic words was provided in the number of academic word families appearing in the textbooks, their frequency distribution, and the word in-depth knowledge. The pedagogical implications were as follows: the occurrences of new words in the texts could be adjusted and controlled so as to ensure learners' learning and use of them, and more attention should be paid to the collocation of words in the textbook designing process, which is vital to realize contextual richness and is conducive to acquire vocabulary.

**Keywords:** Academic Word List (AWL), textbook, vocabulary evaluation, learning opportunity, word frequency

## 1. Introduction

Academic words, called as semi-technical vocabulary or sub-technical vocabulary (Ming-Tzu & Nation, 2004), are described as those which are excluded by West's (1953) most frequent 2000 English words but highly salient in academic texts, and of which are supportive rather than central role to the topics of the texts where they occur (Coxhead, 2000). Some words in the Academic Word List (AWL) are not included in the most frequent 2 000 English words although they are highly salient in academic texts (Coxhead, 2000). Mastering them is a great help for EFL college students to develop their English academic paper reading and writing capacities, for students, especially non-native speakers, are generally not as conversant with academic vocabulary as they are with general English words and technical vocabulary in their own specialty fields (Worthington & Nation, 1996; Xue & Nation, 1984).

Textbook is important in the language learning and teaching. Richards (2001) regards it as a key component in language class and states that "much of the language teaching that occurs around the world today could not take place without the extensive use of commercial textbooks". Mares (2003) posits that textbooks have a significant role on language classes and teaching goes on within them. Reda (2002) found basic English vocabulary is among widely used English textbooks. This empirical study is designed to investigate academic words in a series of the New Standard of College English (NSCE, Henceafter), which is widely used among Chinese colleges and universities and learning opportunities of academic words in the vocabulary input.

### 1.1 Academic Word List (AWL)

Academic words, called as semi-technical vocabulary or sub-technical vocabulary (Ming-Tzu & Nation, 2004), are described as those which are excluded by West's (1953) most frequent 2000 English words but highly salient in academic texts, of supportive rather than central role to the topics of the texts where they occur (Coxhead, 2000). They pose a great number of difficulties for EFL learners (Cohen et al., 1988) because academic vocabularies appear with lower frequency than general-service vocabulary items do (Xue & Nation, 1984). It is found in several scholars' views (Worthington & Nation, 1996; Xue & Nation, 1984) that students, especially

non-native speakers, are generally not as conversant with it as they are with general English words and technical vocabulary in their own specialty fields. But in order to fulfill vocabulary learning and teaching goals in English for Academic Purposes (EAP) programs, one of the most challenging aspects is making principled decisions about what kind of words are worth attention during precious class and independent study time.

In order to identify the most useful academic vocabularies, various word lists have been compiled either by hand or by computer. There are lists on corpora and identified words that appeared across a range of texts (Campion & Elley, 1971; Praninskas, 1972 cited in Coxhead, 2000) and lists by tracking student annotations above words in course books (Lynn, 1973; Ghadessy, 1979 cited in Coxhead, 2000). The above four studies were conducted under no assistance of computers. Editing and combining the four lists mentioned above, Xue and Nation (1984) created the University Word List (UWL). Nevertheless, as a combination and synthesis of the four different studies, it is short of consistent principles of selection and many weaknesses in the previous discoveries are inherited. The corpora on which the researches have been based are small in volume and do not include a wide and representative range of topics.

### *1.2 Textbook Vocabulary Evaluation*

Understanding that textbook is a container of vocabulary input as well as an important resource for language learning and teaching and identifying analysis of language teaching material, textbooks included, is a necessary step for the improvement of English education. Since the 1980s, as with the re-prevalence of cognitive linguistics and rapid growth of corpus linguistics, assessment of language materials, especially textbooks, has attracted the attention of language researchers and instructors. Existing researches on this topic are different as regards languages involved, goal and scope, but in order to classify them in a simple way, we divide between non-corpus/qualitative studies and corpus-based / quantitative studies.

Martínez's (1999) discovery shows that designers of textbooks seldom consider research findings of language teaching and learning via exploring the impacts of language learning theories on the design of vocabulary activities. Investigating whether a type of EFL textbooks used in some Spanish schools reach the standards advocated by the European Framework of Reference, Ojeda (2006) surveys general tendencies concerning vocabulary relevant to social lives in textbooks. In her comparative analysis of the textbooks' vocabulary input from education level at college and at primary middle school respectively, she finds that English teaching textbooks excessively use the words associated with material things and inappropriately emphasize social success connected to money.

Research on the lexical profile of EFL textbooks (Miranda, 1990; Matsuoka & Hirsh, 2010), a survey of various types of input materials including graded readers, writing samples of non-native speakers, novels, newspaper text, etc. (Nation & Wang 1999), a study of the vocabulary load of textbooks to see if it can reach a certain standard (O'Loughlin, 2012) and a comparison of textbooks and authentic corpus data (Ljung, 1991).

Miranda (1990) investigates the vocabulary input in textbooks prescribed for students of secondary middle school; Matsuoka and Hirsh (2010) investigate the vocabulary learning opportunities in an ELT course book designed for upper-intermediate learners, suggesting that the text would provide opportunities to deepen knowledge of the second 1,000 most frequent words in English and a context for pre-teaching of academic words; the study of Nation and Wang (1999) tries to make sure whether their graded readers sample provides good conditions for intensifying vocabulary learning; Ljung (1991) presents evidence of the overuse of concrete words to the detriment of abstract ones, as well as a poor representation of words which are useful in the establishment of communicative interaction and social relationships. O'Loughlin (2012) analyzes input from three levels of the course book series New English File using the computer program VocabProfile, and results indicate that learners who complete three course book levels will receive exposure to fewer than the first 1,500 most frequent words in English.

However, despite the variety of perspectives, the conclusions reached by these studies are basically of similar nature because they point to the shortage of systematic standards on three dimensions: 1) the vocabulary load to be included in textbooks; 2) the frequency distribution of vocabulary; and 3) how the word in-depth knowledge should be presented. However, they seldom investigate vocabulary input of a certain kind of words in textbooks. One aim of the present study is to evaluate academic words in a series of college EFL textbooks according to the above mentioned three criteria in order to ascertain whether there is a systematic or a disorder approach in the vocabulary input included in textbooks of academic words from the AWL.

Altogether, the above content reviews the object of the present study—words from the AWL, two aspects of the research field the textbook vocabulary evaluation.

### 1.3 Research Questions

Using academic words as links, this research aims to explore learning opportunities of academic words provided by *NSCE*. To be specific, it tentatively answers the following questions:

- 1) What learning opportunities of academic words are provided by the *NSCE* in terms of the vocabulary load?
- 2) What learning opportunities of academic words are provided by the *NSCE* in terms of the frequency distribution?
- 3) What learning opportunities of academic words are provided by the *NSCE* in terms of the word in-depth knowledge?

## 2. Methodology

### 2.1 Materials

The *New Standard College English (Integrated Course)*, a series of textbooks planned for undergraduates excluding English majors, is selected for analysis and research because at least in the aspect of vocabulary, it claims that lexis is guided by the wordlist in the *College English Curriculum Requirements* (2007) and supported by the *Macmillan English Dictionary for Advanced Learners* (2nd ed., 2007). This series of textbooks asserts to comply with laws of language learning, emphasize coverage, re-occurrence and ingenious use and help students differentiate word categories and master words' meaning and usage. Simultaneously, they are each equipped with vocabulary booklets, providing all new words, phrases, their Chinese paraphrases and exemplar sentences for key words and phrases. Various ways of vocabulary presenting are effectively combined to expedite texts reading and extracurricular vocabulary learning, improving students' learning efficiency. Each book has 10 units containing the following basic patterns: starting points, active reading (1), talking points (1), active reading (2), talking points (2), language in use, reading across cultures, guided writing, unit task, unit file. All these pages of the four books are transferred to plain text to build a mini-corpus.

### 2.2 Research Instruments

Instruments used in the present study are various and diversified, and they are divided into two types as follows:

#### 2.2.1 Computer Programs

The analysis of this study involves the use of the following computer programs or software: ABBYY FineReader 10.0 Sprint is used for scanning textbooks into plain texts, expediting further computer operations. The on-line software VocabProfile (Cobb, 2009) is used for analyzing distribution features of academic words. Wordsmith 4.0 is used for counting frequencies of academic words, searching their collocations and locating their positions in four textbooks. And SPSS 17.0 is the selected statistical software to process random sampling, statistics computing, correlation & regression analysis and chart making.

#### 2.2.2 Contextual Richness Scale

Levels of contextual richness invented by Beck et al. (1983) (cited in Joe, 2010) are used for measuring the level of contextual support for academic word items in course book texts. The specific indices are as follows:

Table 1. Levels of contextual richness indice

Score	Level	Interpretation
1	Verbatim context	Repeated exposure to the same word forms, collocations or sentences through reading or listening (i.e., no new contextual information added).
2	Nonspecific context	The context does not direct learners to understand a precise or general word meaning (e.g., "What is trigger?")
3	General context	The context provides clues about the semantic field or general category but not sufficiently to define precise properties of the word.
4	Specific context	The context directs learners to a specific meaning that can easily be inferred.

### 2.3 Procedure

- 1) Scan all the pages of the *NSCE* (Book 1~4) into computer and save as plain text for further operation.
- 2) Four converted text profiles of the *NSCE* (Book 1~4) are put into VocabProfile for analyzing, whose working principles of textual words are as follows: it divides all words into four categories according to their frequency, namely, 1) the most frequent 1000 English words (K1); 2) the second most frequent 1000 English words (K2); 3) academic words from the AWL; 4) off-list words.
- 3) Use vocabulary analysis software Wordsmith 4.0 to count the occurrences of academic words. First the headwords of the word families in the Academic Word List, then all the inflexions and derivations. The latter are added up to the former, for example, the frequency of the word family *achieve* is 55, including *achieve* 19 occurrence, *achiever* 3, *achievements* 4, *achievement* 5, *achieves* 5, *achieving* 6 and *achieved* 13.
- 4) A case study is conducted on word in-depth knowledge investigation of 15 items which occur at least 1 time in the text to examine the type(s) of repetitions a learner will meet in terms of word form and collocations. In order to randomly choose 15 items, the words are sorted by increasing frequency of occurrence. The 15 words randomly selected are as follows: *manipulate, denote, accommodate, isolate, abstract, restore, overall, promote, consult, conflict, professional, academy, converse, aware, and define*.

### 3. Results and Discussions

This section is mainly about the description of internal natures of academic words in the textbook, including 1) size and proportion of academic words in the lexical frequency profile of the course books, 2) frequency distribution of academic words, and comparison between frequency of 570 academic words and its corresponding rating values as defined before and 3) word in-depth knowledge embodied by a case study of 15 randomly sampled academic words.

#### 3.1 Size and Proportion

The results of the four converted text profiles are shown in Table 2 and 3:

Table 2. Total number of tokens and percentage of tokens of different levels in four books

	1st	2nd	3rd	4th	average
Total	57986	61874	77438	79846	69286
K1 Words (1-1000)	82.92%	82.11%	82.23%	81.66%	82.23%
K2 Words (1001-2000)	6.83%	6.42%	5.67%	6.15%	6.27%
K1+K2	89.75%	88.53%	87.90%	87.81%	88.50%
AWL Words (academic)	3.36%	3.90%	3.29%	4.14%	3.67%
Off-List Words	6.88%	7.56%	8.82%	8.05%	7.83%

Table 3. Percentage of different levels of word families in four books

Word Families*	1st		2nd		3rd		4th		
K1 Words (1-1000):	965	892	92.43%	895	92.74%	939	97.31%	931	96.47%
K2 Words (1001-2000):	987	571	57.85%	556	56.33%	651	65.96%	665	67.38%
K1+K2	1952	1463	74.95%	1451	74.33%	1590	81.45%	1596	81.76%
AWL Words (academic):	570	311	54.56%	350	61.40%	380	66.67%	426	74.74%
Total	2522	1774	70.34%	1801	71.41%	1970	78.11%	2022	80.17%

Note: \* Set value in VocabProfile. For example, the number 965 means the most frequent English words come from 965 word families.

Although the above is a partial result acquired by analyzing tokens and word families, the data are quite revealing, and able to reflect the lexical frequency profile of a series of four books. From the perspective of vocabulary size (token), Table 2 presents an increasing tendency of total token amount, from 57,986 in the 1st book to 79,846 in the 4th one, up 37.69%. In terms of different levels, K1 words, K2 words, AWL words and Off-list words in the four course books have a steady distribution, thus the average of four books can guarantee their lexical frequency profile, as Graph 4.1 indicates. 88.50% words of the textbooks are from the most frequent 2000 English words, 3.67% from academic words, and another 7.83% are not in K1, K2 and AWL, belonging to low frequency words. But we can still inspect differences performed by K1, K2, AWL and Off-list words among four books: the proportion of K1+K2 (the most frequent 2000 English words) distribution has seen an obviously descending tendency, from 89.75% of all tokens in Book 1 down to 87.81% of that of Book 4; however, the proportion of academic words and off-list words gradually increase, from 3.36% and 6.88% up to 4.14% and 8.05% respectively, implying a gradual raise of vocabulary difficulty in textbooks while keeping the volume of basic words.

From the perspective of word families, Table 3 indicates that words in the textbook have a significantly rising coverage of all K1 words, K2 words and AWL words, which have 965, 987 and 570 word families respectively. The detailed coverage of four books is in Table 2, holistically in rising tendency; among four books the coverage of K1 words keeps in a high and steady level (all above 92%) but the coverage amplitude of K2 words and AWL words are as high as about 10% and 20%, which implies an escalation of textbook difficulty. In other words, we can explain the result as when textbook users finish learning the set, they can contact 939, 665 and 426 K1 words, K2 words and AWL words respectively with the holistic coverage reaching 80.17%.

From analysis of lexical frequency profile of this series of textbooks, we observe an evident increase of the size and proportion of academic words.

### 3.2 Frequency Distribution

The above presents the number and coverage academic words occupied in textbooks, and the following will bring the result of their frequency distribution counted by Wordsmith 4.0 and SPSS 17.0.

From the frequency distribution of academic words, a conclusion is easily drawn that 501 academic word families from the AWL appear in the textbooks and the highest frequency distribution is concentrated in 1~5 and 6~10 occurrence (there are 117 and 121 academic words falling in the two sections, occupying 41.8% (53.9% minus 12.1%) of the total 570 word families), and then the tendency is descending drastically. The repetition rate of academic words is considerable because those occurring twice or above occupy 85.6%. A phenomenon to note is that there are still 12.1% word families never appearing in any of the four textbooks, so that the course books are able to cover 87.9% academic words.

Whether or not the frequency distribution of academic words in the textbooks is congruent with that in the

Coxhead's (2000) Academic Corpus is inspected by a correlation analysis below. Therefore, we have two columns of data: the one is frequency and the other rating value as mentioned in Section 3.5 of the last chapter with both of them 570 items, and a Pearson correlation coefficient is calculated as follows:

Table 4. Pearson correlation between frequency and rating

		Frequency	Rating
	Pearson Correlation	1	.345**
Frequency	Sig. (2-tailed)		.000
	N	570	570
	Pearson Correlation	.345**	1
Rating	Sig. (2-tailed)	.000	
	N	570	570

\*\*. Correlation is significant at the 0.01 level (2-tailed).

$r = .345$  ( $p < .01$ ) is not a high one, which can be expounded as the compilation of academic words in the textbooks is not completely unanimous with their general frequency in ordinary texts. Maybe it is constrained by topics and contents editors have designed or this factor is totally ignored and not considered at all.

### 3.3 Word In-Depth Knowledge

Based on data of frequency distribution of academic words, the current case study examines the details of repetitions of a random sample of these words to further inspect the range for deepening knowledge of such words. A case study of 15 randomly selected words is conducted to see how many aspects of word in-depth knowledge are presented to the reader with each occurrence of the chosen words. Table 5 summarizes the frequency, range of the case study words and what aspects of word in-depth knowledge are presented in the text for each chosen word.

Table 5. Case study details

Selected headwords (frequency in parentheses)	Number of chapters they occur in (frequency in parentheses if more than 1 per chapter listed)	Part of speech (frequency in parentheses)	Number of inflections and derivations	Number of collocations**
Manipulate (3)	R: 1 = b4u9 ( $\times 3$ )*	Verb (3)	1 inflection: Manipulate = 3	0
Denote (4)	R: 3 = b4u5 ( $\times 2$ ) + b4u8 + b4g	Verb (3) Noun (1)	1 inflection: Denote = 3 1 derivation: Denotation = 1	0
Accommodate (5)	R: 4 = b1u8 + b2u1 + b4u1 + b4u7 ( $\times 2$ )	Noun (5)	1 derivation: Accommodation = 5	Accommodation = 5
Isolate (6)	R: 4 = b2u6 + b4u5 + b4u9 ( $\times 3$ ) + b4g	Verb (1) Adjective (5)	1 inflection: Isolated = 6	Isolated = 5

Abstract (8)	R: 5 = b2u7 ( $\times 4$ ) + b4u2 + b4u4 + Adjective (8) b4u6 + b4u8	Adjective (8)	1 inflection: Abstract = 8	Abstract = 8
Restore (8)	R: 5 = b1g + b1u10 ( $\times 3$ ) + b4u3 + b4u6 ( $\times 2$ ) + b4u7	Verb (7) Noun (1)	3 inflections: Restore = 3 Restored = 3 Restoring = 1 1 derivation: Restoration = 1	0
Overall (10)	R: 6 = b1u8 + b2u2 + b2u9 + b2g + b3u4 ( $\times 3$ ) + b4u8 ( $\times 3$ )	Adjective (8) Adverb (2)	1 inflection: Overall = 10	Overall = 10
Promote (13)	R: 9 = b2u9 + b3u5 + b3u6 ( $\times 2$ ) + b3g + b4u1 ( $\times 2$ ) + b4u3 + b4u5 + b4u6 ( $\times 2$ ) + b4u9 ( $\times 2$ )	Verb (8) Noun (5)	2 inflections: Promote = 5 Promoted = 3 1 derivation: Promotion = 5	Promote = 5 Promotion = 5
Consult (16)	R: 10 = b1u3 ( $\times 2$ ) + b2u2 + b2u3 ( $\times 4$ ) + b2u8 + b2g + b3u7 + b4u1 ( $\times 2$ ) + b4u4 ( $\times 2$ ) + b4u10 + b4g	Verb (7) Noun (7) Adjective (2)	4 inflections: Consult = 2 Consults = 1 Consulting = 4 Consulted = 2 4 derivations: Consultancy = 3 Consultant = 1 Consultation = 1 Consultations = 2	0
Conflict (18)	R: 9 = b1u4 + b2u3 + b2u5 ( $\times 8$ ) + b2u7 + b2u9 + b3u4 + b3u5 ( $\times 2$ ) + b3u9 + b4u1 ( $\times 2$ )	Noun (17) Adjective (1)	3 inflections: Conflict = 10 Conflicting = 1 Conflicts = 7	Conflict = 9 Conflicts = 7
Professional (21)	R: 12 = b1u5 + b1u6 + b1u9 ( $\times 2$ ) + b2u3 + b2u4 + b2u6 ( $\times 4$ ) + b2u8 ( $\times 3$ ) + b2g + b3u3 + b3u7 + b4u4 + b4u5 ( $\times 4$ )	Noun (2) Adjective (17) Adverb (2)	3 inflections: Professional = 18 Professionally = 2 Professionals = 1	Professional = 18
Academy (27)	R: 16 = b1u1 + b1u3 + b1u7 + b2u1 ( $\times 4$ ) + b2u9 + b3u5 + b3u7 ( $\times 3$ ) + b3u9 + b3u10 + b4u1 ( $\times 2$ ) + b4u5 ( $\times 4$ ) + b4u6 ( $\times 2$ ) +	Noun (5) Adjective (22)	1 inflection: Academy = 3 2 derivations: Academic = 22 Academics = 2	Academic = 22

	b4u7 + b4u8 ( $\times 2$ ) + b4u9 + b4u10	
Converse (31)	R: 15 = b1u9 + b1u10 + b2u1 + b2u4 + b3u3 + b3u4 ( $\times 2$ ) + b3u5 + b3u10 ( $\times 3$ ) + b3g + b4u1 ( $\times 2$ ) + b4u2 + b4u3 + b4u5 ( $\times 12$ ) + b4u7 ( $\times 2$ ) + b4u9	1 inflection: Converse = 4 Conversations = 5 2 derivations: Conversation = Conversation = 22 Conversations = 5
Aware (57)	R: 24 = b1u1 + b1u3 + b1u4 ( $\times 3$ ) + b1u7 + b1u8 ( $\times 5$ ) + b1u10 ( $\times 9$ ) + b1g + b2u1 + b2u2 + b2u3 ( $\times 2$ ) + b2u6 + b2u7 ( $\times 8$ ) + b2u9 ( $\times 3$ ) + b3u1 + b3u3 ( $\times 5$ ) + b3u4 ( $\times 2$ ) + b3u5 ( $\times 2$ ) + b3u8 ( $\times 2$ ) + b3u10 + b4u1 ( $\times 2$ ) + b4u2 ( $\times 2$ ) + b4u7 + b4u8 + b4u9	Noun (28) Adjective (29)
Define (135)	R: 42 = b1u1 + b2u2 ( $\times 3$ ) + b1u3 ( $\times 2$ ) + b1u4 + b1u5 ( $\times 2$ ) + b1u6 ( $\times 3$ ) + b1u7 ( $\times 4$ ) + b1u8 ( $\times 3$ ) + b1u9 ( $\times 2$ ) + b1u10 ( $\times 3$ ) + b1g + b2u1 ( $\times 11$ ) + b2u2 ( $\times 3$ ) + b2u3 ( $\times 2$ ) + b2u4 ( $\times 3$ ) + b2u5 ( $\times 3$ ) + b2u6 ( $\times 2$ ) + b2u7 ( $\times 3$ ) + b2u8 ( $\times 2$ ) + b2u9 ( $\times 5$ ) + b2u10 ( $\times 5$ ) + b2g + b3u1 ( $\times 2$ ) + b3u2 ( $\times 2$ ) + b3u3 ( $\times 2$ ) + b3u4 ( $\times 3$ ) + b3u5 ( $\times 2$ ) + b3u6 ( $\times 9$ ) + b3u7 ( $\times 6$ ) + b3u8 ( $\times 2$ ) + b3u9 ( $\times 2$ ) + b3u10 ( $\times 2$ ) + b4u1 ( $\times 2$ ) + b4u2 ( $\times 5$ ) + b4u3 ( $\times 3$ ) + b3u4 ( $\times 2$ ) + b4u5 ( $\times 4$ ) + b4u6 ( $\times 4$ ) + b4u7 ( $\times 4$ ) + b4u8 ( $\times 8$ ) + b4u9 ( $\times 3$ ) + b4u10 ( $\times 3$ )	1 inflection: Aware = 27 Aware = 27 2 derivations: Awareness = 28 Awareness = 28 Unaware = 2
		4 inflections: Define = 22 Defined = 10 Defined = 10 Definition = 18 Defines = 3 Define = 22 Defining = 3 Definitions = 79 2 derivations: Definition = 18 Definitions = 79 Definitions = 79

Note: \* b is short for book, u for unit, and g for glossary. \*\* based on data of collocates from Wordsmith 4.0.

Table 6. Scores of levels of contextual richness

Selected headwords	Verbatim context	Nonspecific context	General context	Specific context	scores
	1 (frequency in parentheses)	2 (frequency in parentheses)	3 (frequency in parentheses)	4 (frequency in parentheses)	
manipulate	2 (2)		3 (1)		5
denote	3 (3)			4 (1)	7
accommodate		2 (1)	9 (3)	4 (1)	15
isolate	3 (3)		6 (2)	4 (1)	13
abstract		2 (1)	18 (6)	4 (1)	24
restore	3 (3)		12 (4)	4 (1)	19
overall	3 (3)		21 (7)		24
promote	2 (2)	2 (1)	30 (10)		34
consult	4 (4)	2 (1)	42 (14)		48
conflict			45 (15)	12 (3)	57
professional	1 (1)		60 (20)		61
academy	1 (1)	8 (4)	63 (21)	4 (1)	76
converse	2 (2)	12 (6)	69 (23)		83
aware	7 (7)	2 (1)	141 (47)	8 (2)	158
define	2 (2)	148 (74)	174 (58)	4 (1)	328

Table 7. Pearson correlation between frequency of selected headwords and scores of levels of contextual richness

		Frequency of selected headwords	Scores of levels of contextual richness
Frequency of selected headwords	Pearson Correlation	1	.997**
	Sig. (2-tailed)		.000
	N	15	15
Scores of levels of contextual richness	Pearson Correlation	.997**	1
	Sig. (2-tailed)	.000	
	N	15	15

\*\*. Correlation is significant at the 0.01 level (2-tailed).

Above three tables clearly evidence details of selected academic words and scrutinize natures of them throughout four books. The more frequent a word appears in the textbook, the more possible the word will be in different parts of the textbook such as text body part, supplementary materials, appended exercises, review parts, etc, thus more word in-depth knowledge will be delivered. A general tendency can be concluded that number of chapters the selected words occur in, number of inflections, derivations and collocations, and scores of levels of contextual richness all increase with frequency of selected headwords; they are in positive correlation. For example, Pearson correlation coefficient between frequency of selected headwords and scores of levels of contextual richness is as high as .997, indicating that the more frequent a word is, the higher the levels of contextual richness it gets.

### 3.4 Learning Opportunities of Academic Words Provided in the NSCE

Studying the number of academic words contained in textbooks and their frequency distribution can cast enlightenments on how the learners are exposed to their input at different grade levels (from Book 1 to 4). Furthermore, it can offer instructors and researchers with meritorious information regarding the reasonable arrangement of the number and frequency of academic words contained in textbooks so as to play its effect into full swing. The present study reveals that the analyzed four course books display remarkable disparities of academic words concerning the size and proportion, frequency distribution, aspects of word in-depth knowledge.

Table 8. Distribution statistics of academic words in four books

Academic words	1st	2nd	3rd	4th
Token	1951	2416	2548	3306
Type	535	629	699	863
TTR	27.42%	26.03%	27.43%	26.10%
Word families	311	350	380	426

Firstly, we discuss the learning opportunities of academic words provided by their size and proportion. In Section 3.1, we have a glimpse of the size and proportion of academic words that are included in the textbook lexical frequency profile. As described in detail in Table 8, it can be seen that volumetric gaps of tokens, types and word families exist among books with the one between the 2nd and 3rd volume a little smaller. The constant type-token ratio, remaining at 26%~27%, suggests every one type has averagely been embodied about five tokens. As regards the word families each textbook covers, the percentage rises from 54.56% to 74.74%, a nearly one fifth increase. Such differences and changes may have influences on students since they will receive an escalating amount of academic words' input provision when following this series of materials. Although only 311 academic word families (54.56%) are covered at the beginning, the learners will be ensured exposure of 501 ones (87.9%) (four books combined together not the fourth textbook alone) in the end.

Table 9. Occurrence of academic word families in four books

Occurrences	Words from the AWL (501 word families found)
once	2.3%
6 times or more	17.7%
10 times or more	24.2%
More than 20 times	25.4%

Secondly, attention goes to frequency distribution. Table 9 gives us information about it. Despite some scholars' thought that little correlation is found between the number of frequencies and vocabulary acquisition and that the type of text where the word appears and is encountered may have a greater influence than the number of exposures (Brown, 1993), many literatures in second language acquisition record the contribution of encounters for vocabulary acquisition even though lacking an unanimous agreement on the accurate times (Krachroo, 1962; Jenkins & Dixon, 1987; Bunker, 1988). Of 501 academic word families occurring in the textbooks, only 2.3% have occurred once, which are behalf, complement, federal, framework, ideology, implicit, insert, inspect, intermediate, minimal, simulate, somewhat, and underlie. Nagy and Anderson (1984) and Nagy and Herman (1987) have maintained that only one encounter of a new word is significant and may result in partial acquisition of its meaning, but the majority of academic words are intrinsically abstract, and only one encounter is far from sufficient for incidental acquisition, as is verified by subsequent test that averagely one input frequency only leads to 21.81% academic acquisition rate. While those occurring above 6 times are accounting for 67.4% overall, denoting that more than half academic words in textbooks are well repeated to the extent conducive to acquisition because Rott (1999) finds six encounters of a new word produce significantly more vocabulary knowledge growth and it is a critical point for second language vocabulary acquisition. Word repetition is also the index gauging text comprehension difficulty. Bunker (1988) finds a strong correlation between the percentages of words repeated and reading difficulty in L2, showing that those texts that repeated about 33

percent of their words more than five times were easier than those that repeated 20, 19, or 14 percent of their words at least five times. He also found that as the number of percent of words repeated at least five times in a text increased, the difficulty of learning words decreased. A high proportion of repetition for 67.4% academic words in this series of course books is a witness for its low text difficulty, coincidentally, as textbooks they should not have been too difficult.

Last but not least, the result of 15-word case study is explicated to see how many aspects of word in-depth knowledge are presented. They are sampled to reflect the details and essences of word occurrence across texts, revealing how the aspects of word knowledge are deepened within academic words. We shall scrutinize Table 5 column by column: 1) the item number of chapters they occur in is the range one word family distributed in the whole texts, which stands for spaced repetition (i.e., the spreading out of the repetitions of a word throughout the text) rather than massed repetition (i.e., the concentration of the repetitions of a word in only one part of the text). Bloom and Shuell (1981), Dempster (1987) and Baddeley (1990) have found that patterns of spaced repetition contribute to building and storing memory, leading to longer and more enduring word retention with less attrition. Nation (2001) holds that learning from repetition not only concerns the numbers of repetition but also its spacing; 2) the number of appearing parts of speech is highly relevant to word frequency and more variables predict more contacts, thus bring learners with more opportunities to master their semantic dimensions and building relations between meanings; 3) compatible to the case study of Matsuoka and Hirsh (2010), more inflections are found than derivation with a ratio of 27: 16, partly because all the inflections and derivations of academic word families are too abstract and uncommon. For instance, the word family abstract has as many as eight inflections and derivations altogether, but only one inflection abstract appears. Others, in Sublist 6 of the AWL, such as abstractness, abstraction, and abstractions serving as noun, abstracts, abstracted, and abstracting serving as noun or verb, and abstractly serving as adverb are relatively infrequent even in native speakers' usage, let along in this series of EFL college course books; 4) collocation is the specific way to deepen word knowledge, converting an unknown or partially known word to the fullest possible. The word family having most collocations goes to define, also the highest occurring one of the 15 items, whose collocations includes two inflections define and defined having 22 and 10, two derivations definition and definitions having 18 and 79.

An example in point of how data are collected on word in-depth knowledge in this case study can be stood by the entry for *consult* in Table 5. The academic word family *consult* occurs 16 times in 10 different chapters, 7 times as a verb (*consult, consults, consulting, consulted*), 7 times as a noun (*consultancy, consultation, consultations*), and twice as an adjective (*consulting, consultant*). 4 inflections (*consult, consults, consulting, consulted*) of *consult* are demonstrated in several tenses (simple present, third person simple present, present progressive, present perfect, and present passive). 4 derivations (*consultancy, consultant, consultation, consultations*) appearing either individually in the glossary or contextually in the articles or exercises. The multiple inflections and derivations expose learners to multivariate uses and semantic meanings, delivering them a great opportunity of learning academic words. Although data from Wordsmith 4.0 indicate that no collocation of the academic word family *consult* is found, judging from concordances, we extract *consult sb/sth, run a consultancy*, which is not as abundant as other aspects of word in-depth knowledge.

#### 4. Conclusion

Overall, the number of academic word families appearing in the textbooks, their frequency distribution, and the word in-depth knowledge revealed by the case study all suggest favorable opportunities be provided, which are featured by high degree of coverage of all academic word families, sufficient and spaced repetition, various and abundant parts of speech, inflections and derivations. Insufficiently, learning opportunities of academic words provided by collocations are not as profuse and lavish as other aspects of word in-depth knowledge.

Theoretical implications drawn from this study sheds light on our understanding of the applicability of the AWL to the textbook. Its founder Coxhead (2000) claims that the 570 word families from the AWL constitute a specialized vocabulary with good coverage of academic texts regardless of the subject area, accounting for 10% of the total tokens in the Academic Corpus. Compared with the findings in this study, where words from the AWL have a coverage from 3.36% to 4.14%, it is found that this series of college English textbooks is not EAP oriented because the AWL coverage is much lower than that of the Academic Corpus, although favorable provisions of learning opportunities of academic words exist.

The results obtained from this study have important pedagogical implications for learners, teachers, materials writers and researchers. Learners and teachers should encourage students to make full use of textbooks. As revealed previously, students surveyed in this study have not always done exercises appended to the texts, where a great number of words required to master repeat, academic words included. Since textbooks can provide such a

favorable opportunity for vocabulary learning and students are required to have English classes based on textbooks, they might just as well make the best of textbooks first to partially fulfill their learning goals efficiently. Materials writers and researchers should avail of plenty of vocabulary researches such as Vocabprofile (Cobb, 2009) in the process of putting materials together and compiling textbooks so as to strengthen the function of textbook as an important source for vocabulary input in accordance with the facilitative role of the input frequency effect, maximizing the most frequent words or those required to be mastered by the syllabus and minimizing low frequency vocabulary.

## References

- Asher, R. E. (1994). *The Encyclopedia of Language and Linguistics*. Exeter: Pergamon Press.
- Baddeley, A. (1990). *Human Memory*. London: Lawrence Erlbaum Associates.
- Bloom, K. C., & Shuell, T. J. (1981). Effects of massed and distributed practice on the learning and retention of second-language vocabulary. *Journal of Educational Research*, 74, 245-248.
- Brown, C. (1993). Factors Affecting the Acquisition of Vocabulary: Frequency and Saliency of Words. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second Language Reading and Vocabulary Learning* (pp. 263-286). Norwood, NJ: Ablex.
- Bunker, E. L. (1988). *Towards comprehensive guidelines for the preparation of material in English for EFL adults with limited education* (M.A. Thesis, Linguistics Department, Brigham Young University. Provo, Utah).
- Chen, Q., & Ge, G. (2007). A corpus-based lexical study on frequency and distribution of Coxhead's AWL word families in medical research articles (RAs). *English for Specific Purpose*, 26(4), 502-514. <http://dx.doi.org/10.1016/j.esp.2007.04.003>
- Cobb, T. (2009). *Web VocabProfile* (Version: VP English v.3). Retrieved from <http://www.lextutor.ca/vp>
- Cohen, A., Glasman, H., Rosenbaum-Cohen, P. R., Ferrara, J., & Fine, J. (1988). Reading English for specialised purposes: Discourse analysis and the use of standard informants. In P. Carrell, J. Deline, & D. Eskey (Eds.), *Interactive Approaches to Second Language Reading* (pp. 152-167). Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9781139524513.017>
- Coxhead, A. (2000). A New Academic Word List. *TESOL Quarterly*, 34(2), 213-238. <http://dx.doi.org/10.2307/3587951>
- Dempster, F. N. (1987). Effects of variable encoding and spaced presentation on vocabulary learning. *Journal of Educational Psychology*, 79, 162-170. <http://dx.doi.org/10.1037/0022-0663.79.2.162>
- Harmer, J. (1991). *The Practice of English Language Teaching*. London: Longman.
- Jenkins, J., & Dixon, R. (1983). Vocabulary learning. *Contemporary Educational Psychology*, 8, 237-260. [http://dx.doi.org/10.1016/0361-476X\(83\)90016-4](http://dx.doi.org/10.1016/0361-476X(83)90016-4)
- Joe, A. (2010). The quality and frequency of encounters with vocabulary in an English for Academic Purpose programme. *Reading in a Foreign Language*, 22(1), 117-138.
- Joe, A., Nation, I. S. P., & Newton, J. (1996). Vocabulary learning and speaking activities. *English Teaching Forum*, 34(1), 2-7.
- Kachroo, J. N. (1962). Report on an investigation into the teaching of vocabulary in the first year of English. *Bulletin of the Central Institute of English*, 2, 67-72.
- Laufer, B., & Nation, I. S. P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16(3), 307-322.
- Liu, H. (2013). A Literature Review on College English Teaching Materials Study in the Past 20 Years. *Computer-assisted Foreign Language Education*, 150, 66-71.
- Ljung, M. (1991). Swedish TEFL meets reality. In S. Johansson, & A. B. Stenström (Eds.), *Computer Corpora: Selected Papers and Research Guide* (pp. 245-256). Berlin: Mouton de Gruyter.
- Mares, C. (2003). Writing a coursebook. In B. Tomlinson (Ed.), *Developing Materials for Language Teaching* (pp. 130-140). New York: Continuum.
- Martínez, B. (1999). *The influence of language learning theories on the design of vocabulary activities in L2 (English) course books*. University of La Rioja: Departamento de Filologías Modernas: Small thesis.

Unpublished Manuscript.

- Matsuoka, W., & Hirsh, D. (2010). Vocabulary learning through reading: Does an ELT course book provide good opportunities? *Reading in a Foreign Language*, 22(1), 56-70.
- Ming-Tzu, K. W., & Nation, P. (2004). Word Meaning in Academic English: Homography in the Academic Word List. *Applied Linguistics*, 25(3), 291-314. <http://dx.doi.org/10.1093/aplin/25.3.291>
- Miranda, A. (1990). The vocabulary of the English Coursebooks: An Analysis. *Revista Espa-ola de Lingüística Aplicada*, 6, 11-117.
- Mudraya, O. (2006). Engineering English: A lexical Frequency instructional model. *English for Specific Purposes*, 25(2), 235-256. <http://dx.doi.org/10.1016/j.esp.2005.05.002>
- Nagy, W. E. (2005). Why vocabulary instruction needs to be long-term and comprehensive. In E. H. Hiebert, & M. L. Kamil (Eds.), *Teaching and learning vocabulary: Bringing research to practice*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Nagy, W. E., & Anderson, R. C. (1984). How many words are there in printed school English? *Reading Research Quarterly*, 19, 304-330. <http://dx.doi.org/10.2307/747823>
- Nagy, W. E., Anderson, R. C., & Herman, P. (1987). Learning word meanings from context during normal Reading. *American Educational Research Journal*, 24, 237-270. <http://dx.doi.org/10.3102/00028312024002237>
- Nation, P. (1990). *Teaching and Learning Vocabulary*. New York: Newbury House/Harper Row.
- Nation, P. (2001). *Learning Vocabulary in Another language*. Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9781139524759>
- Nation, P. (2008). *Teaching Vocabulary: Strategies and Techniques*. Boston: Heinle Cengage Learning.
- Nation, P., & Coady, J. (1988). Vocabulary and Reading. In R. Carter, & M. McCarthy (Eds.), *Vocabulary and language teaching* (pp. 97-110). NY: Longman Inc.
- Nation, P., & Wang, K. (1999). Graded readers and vocabulary. *Reading in a Foreign Language*, 12(2), 355-80.
- O'Loughlin, R. (2012). Tuning In to Vocabulary Frequency in Coursebooks. *RELC Journal*, 43(2), 255-269. <http://dx.doi.org/10.1177/0033688212450640>
- Ojeda, J. (2006). Vocabulary of Success in EFL Text Books (pp. 974-983). *Actas VIII Congreso de la Sociedad Espa-ola de la Lengua y la Literatura*. Universidad de la Habana, Cuba.
- Reada, G. (2003). English Coursebooks: Prototype Texts and Basic Vocabulary Norms. *ELT Journal*, 57(3), 260-268.
- Richards, J. (2001). The role of textbooks in a language program. *Guidelines*, 23(2), 12-16.
- Rott, S. (1999). The Effect of Exposure Frequency on Intermediate Language Learners' Incidental Vocabulary Acquisition and Retention through Reading. *Studies in Second Language Acquisition*, 21(4), 589-619. <http://dx.doi.org/10.1017/S0272263199004039>
- Scholfield, P. (1991). Vocabulary rate in course books: Living with an unstable lexical economy. Thessaloniki: Aristotle University.
- Thornbury, S. (2002). *How to Teach Vocabulary*. Harlow: Longman.
- West, M. (1953). *A general service list of English words*. London: Longman.
- Worthington, D., & Nation, P. (1996). Using texts to sequence the introduction of new vocabulary in an EAP course. *RELC Journal*, 27(2), 1-11. <http://dx.doi.org/10.1177/003368829602700201>
- Xue, G., & Nation, P. (1984). A university word list. *Language Learning and Communication*, 3, 215-229.

### **Copyrights**

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).